

What is claimed is:

1. A method for backing up one or more data files, the method comprising:

detecting changed locations in the one or more data files,

at a storage time, storing the contents of the changed locations, and

associating the stored contents with: the storage time, the changed locations, and one or more file identifiers identifying the one or more data files.
2. The method of claim 1, wherein the storage time is based on one or more of: an actual time, a time interval, and an event.
3. The method of claim 1, further comprising:

prior to detecting changed locations, generating a baseline image, where the baseline image includes at least the one or more data files.
4. The method of claim 3, wherein generating a baseline image includes:

generating one or more of: a volume image, a file image, and a snapshot image.
5. The method of claim 1, wherein detecting changed locations includes:

at a time prior to the storage time,

generating a baseline image of at least the one or more data files, and,

using one or more data integrity procedures to generate a summary of the baseline image;

thereafter,

generating a second image of at least the one or more data files, and,

using the one or more data integrity procedures to generate a summary of the second image; and,

based on the baseline summary and the second summary, determining whether the one or more data files includes changed locations.

6. The method of claim 5, wherein the one or more data integrity procedures include one or more of: a cyclic redundancy check procedure and an MD5 procedure.

7. The method of claim 1, wherein detecting changed locations includes dynamically detecting the changed locations in the one or more data files.

8. The method of claim 7, wherein dynamically detecting the changed locations includes:

at a time prior to the storage time, generating a baseline image of at least the one or more data files, and,

thereafter, dynamically detecting the changed locations.

9. The method of claim 8, further comprising:

iteratively returning to dynamically detecting the changed locations to the one or more data files.

10. The method of claim 1, wherein storing the contents includes:

selecting at least one memory to store the contents.

11. The method of claim 10, wherein selecting at least one memory includes:

selecting the at least one memory to be distinct from a previously selected memory associated with a prior storage time.

12. The method of claim 10, further comprising:

iteratively returning to detecting changed locations.

13. The method of claim 1, wherein associating includes:

generating one or more indexes to associate: the stored contents, the respective storage times, the respective changed locations, and the respective one or more file identifiers.

14. The method of claim 13, wherein the one or more indexes include:

a first index to the changed locations based on the one or more file identifiers, and

a second index to the stored contents based on the changed locations.

15. The method of claim 1, further comprising:

using the stored contents to create a version of a selected one of the one or more data files.

16. The method of claim 15, wherein using the stored contents to create a version includes:

for each of one or more storage times associated with the version, and based on the file identifier associated with the selected data file: querying one or more indexes that associate the stored contents, the respective storage times, the respective changed locations, and the respective one or more file identifiers, to identify stored contents and respective changed locations associated the selected data file, and

combining the identified stored contents with data from a baseline image associated with the selected data file.

17. The method of claim 16, wherein querying includes:

determining that the changed locations are the same for two or more different storage times, and,

identifying the stored contents of the changed locations associated with the latest of the two or more different storage times.

18. The method of claim 1, further comprising:

receiving from a first server a request to create a version of a selected one of the one or more data files, and

based on the request:

for each of one or more storage times associated with the version, and based on the file identifier associated with the selected data file: querying one or more indexes that associate the stored contents, the respective storage times, the respective changed locations, and the respective one or more file identifiers, to identify stored contents and respective changed locations associated the selected data file, and

providing the identified stored contents and respective changed locations to the first server.

19. The method of claim 18, further comprising:

at the first server, combining the identified stored contents with data from a baseline image associated with the selected data file.

20. The method of claim 1, further comprising:

at a coalescence time, coalescing:

two or more stored contents associated with the same file identifier and two or more different storage times,

the respective changed locations associated with the two or more coalesced contents, and

one or more indexes to associate the coalesced contents, the respective coalesced changed locations, the file identifier, and the latest of the two or more different storage times.

21. The method of claim 1, further comprising:

at a coalescence time, coalescing:

two or more stored contents associated with the same file identifier and the same storage time,

the respective changed locations associated with the two or more coalesced contents, and

one or more indexes to associate the coalesced contents, the respective coalesced changed locations, the file identifier, and the same storage time.

22. The method of claim 21, wherein the coalescence time is based on one or more of: an actual time, a time interval, and an event.

23. The method of claim 22, wherein the event includes an event based on an available storage capacity of a storage medium.

24. A method of backing up data files, comprising:

at a first time, storing a baseline image of one or more data files,

detecting changed locations in the one or more data files,

at a storage time, storing the contents of the changed locations, and,

generating one or more indexes to associate the stored contents, the respective storage time, the respective changed locations, and one or more file identifiers to identify the respective one or more data files having the respective changed locations.

25. The method of claim 20, further comprising:

recreating a version of a data file.

26. The method of claim 25, wherein recreating includes:

writing to the version the baseline image associated with the data file,

associating the version with one or more storage times, and

for each of the one or more storage times:

based on querying the one or more indexes, identifying stored contents and respective changed locations associated with the file identifier for the data file, and

writing to the version the stored contents of the respective changed locations.

27. The method of claim 26, wherein recreating further includes:

determining that the changed locations for two or more different storage times are the same, and

wherein writing includes:

writing to the version the stored contents of the changed locations associated with the latest storage time.

28. The method of claim 24, further comprising:

receiving a request from a first server to recreate a version of a data file, and

based on the request:

associating the version with one or more storage times,

for each of the one or more storage times, identifying stored contents and respective changed locations associated with the file identifier for the data file based on querying the one or more indexes, and

providing the stored contents and respective changed locations to the first server.

29. The method of claim 28, further comprising:

associating the version with a version time and one or more storage times.

30. The method of claim 29, further comprising:

at the first server, writing to the version:

the baseline image associated with the data file,

the stored contents of the respective changed locations, and

contents of changed locations in the data file that were dynamically detected between the latest of the one or more storage times and the version time.

31. The method of claim 30, wherein coalescing includes:

iteratively coalescing two or more portions of the stored contents.

32. The method of claim 28, further comprising:

at the first server, writing to the version the baseline image associated with the data file and the stored contents of the respective changed locations.

33. The method of claim 24, further comprising:

at a coalescence time, coalescing:

two or more stored contents associated with the same file identifier and two or more different storage times,

the respective changed locations associated with the two or more coalesced contents, and

the one or more indexes to associate the coalesced contents, the respective coalesced changed locations, the file identifier, and the latest of the two or more different storage times.

34. The method of claim 24, further comprising:

at a coalescence time, coalescing:

two or more stored contents associated with the same file identifier and the same storage time,

the respective changed locations associated with the two or more coalesced contents, and

the one or more indexes to associate the coalesced contents, the respective coalesced changed locations, the file identifier, and the same storage time.

35. The method of claim 34, wherein the coalescence time is based on one or more of: an actual time, a time interval, and an event.

36. The method of claim 35, wherein the event includes an event based on an available storage capacity of a storage medium.

37. A processor program for backing up data files, the processor program being stored on a processor-readable medium and including instructions to cause a processor to:

detect changed locations in one or more data files,

at a storage time, store the contents of the changed locations, and

associate the stored contents with: the storage time, the changed locations, and one or more file identifiers identifying the one or more data files.

38. The processor program of claim 37, wherein the storage time is based on one or more of: an actual time, a time interval, and an event.

39. The processor program of claim 37, wherein the instructions to detect changed locations include instructions to dynamically detect the changed locations in the one or more data files.

40. The processor program of claim 37, wherein the instructions to store the contents include instructions to select at least one memory to store the contents.

41. The processor program of claim 40, wherein the instructions to select at least one memory include instructions to select the at least one memory to be distinct from a previously selected memory associated with a prior storage time.
42. The processor program of claim 40, further comprising instructions to iteratively return to the instructions to detect changed locations.
43. The processor program of claim 37, wherein the instructions to associate includes instructions to generate one or more indexes to associate: the stored contents, the respective storage times, the respective changed locations, and the respective one or more file identifiers.
44. The processor program of claim 43, wherein the one or more indexes include:
- a first index to the changed locations based on the one or more file identifiers, and
 - a second index to the stored contents based on the changed locations.
45. The processor program of claim 37, further comprising instructions to use the stored contents to create a version of a selected one of the one or more data files.
46. The processor program of claim 45, wherein instructions to use the stored contents to create a version include instructions to: for each of one or more storage times associated with the version, and based on the file identifier associated with the selected data file: query one or more indexes that associate the stored contents, the respective storage times, the respective changed locations, and the respective one or more file identifiers, to identify stored contents and respective changed locations associated the selected data file, and
- combine the identified stored contents with data from a baseline image associated with the selected data file.
47. The processor program of claim 37, further comprising instructions to receive from a first server a request to create a version of a selected one of the one or more data files, and
- based on the request:

for each of one or more storage times associated with the version, and based on the file identifier associated with the selected data file: query one or more indexes that associate the stored contents, the respective storage times, the respective changed locations, and the respective one or more file identifiers, to identify stored contents and respective changed locations associated the selected data file, and

provide the identified stored contents and respective changed locations to the first server.

48. The processor program of claim 47, further comprising instructions to, at the first server, combine the identified stored contents with data from a baseline image associated with the selected data file.

49. A system for backing up data files, the system comprising:

one or more servers in communication with the data files, one or more of the servers configured to execute change requests on the data files, and

one or more agents in communication with the one or more servers, the agents configured to:

detect changed locations in the one or more data files,

at a storage time, store the contents of the changed locations, and

associate the stored contents with: the storage time, the changed locations, and one or more file identifiers identifying the one or more data files.

50. The system of claim 49, wherein the storage time is based on one or more of: an actual time, a time interval, and an event.

51. The system of claim 49, wherein the configuration to detect changed locations include a configuration to dynamically detect the changed locations in the one or more data files.

52. The system of claim 49, wherein the configuration to store the contents include a configuration to select at least one memory to store the contents.

53. The system of claim 52, wherein the configuration to select at least one memory include a configuration to select the at least one memory to be distinct from a previously selected memory associated with a prior storage time.

54. The system of claim 52, where the agents are further configured to iteratively return to the detect changed locations.

55. The system of claim 49, wherein the configuration to associate includes a configuration to generate one or more indexes to associate: the stored contents, the respective storage times, the respective changed locations, and the respective one or more file identifiers.

56. The system of claim 55, wherein the one or more indexes include:

a first index to the changed locations based on the one or more file identifiers, and

a second index to the stored contents based on the changed locations.

57. The system of claim 49, further comprising processor instructions to use the stored contents to create a version of a selected one of the one or more data files.

58. The processor program of claim 57, wherein the processor instructions to use the stored contents to create a version include processor instructions to: for each of one or more storage times associated with the version, and based on the file identifier associated with the selected data file: query one or more indexes that associate the stored contents, the respective storage times, the respective changed locations, and the respective one or more file identifiers, to identify stored contents and respective changed locations associated the selected data file, and

combine the identified stored contents with data from a baseline image associated with the selected data file.

59. The system of claim 49, further comprising processor instructions to receive from a first server a request to create a version of a selected one of the one or more data files, and

based on the request:

for each of one or more storage times associated with the version, and based on the file identifier associated with the selected data file: query one or more indexes that associate the stored contents, the respective storage times, the respective changed locations, and the respective one or more file identifiers, to identify stored contents and respective changed locations associated the selected data file, and

provide the identified stored contents and respective changed locations to the first server.

60. The system of claim 59, further comprising processor instructions to, at the first server, combine the identified stored contents with data from a baseline image associated with the selected data file.